DEPT. OF MATH./CMA UNIV. OF OSLO
PURE MATHEMATICS No 13
ISSN 0806–2439 AUGUST 2007

A BAYES FORMULA FOR NON-LINEAR FILTERING WITH GAUSSIAN AND COX NOISE

V. MANDREKAR, T. MEYER-BRANDIS, AND F. PROSKE

ABSTRACT. A Bayes type formula is derived for the non-linear filter where the observation contains both general Gaussian noise as well as Cox noise whose jump intensity depends on the signal. This formula extends the well know Kallianpur-Striebel formula in the classical non-linear filter setting. We also discuss Zakai type equations for both the unnormalized conditional distribution as well as unnormalized conditional density in case the signal is a Markovian jump diffusion.

1. INTRODUCTION

The general filtering setting can be described as follows. Assume a partially observable process $(X,Y) = (X_t, Y_t)_{0 \le t \le T} \in \mathbb{R}^2$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The real valued process X_t stands for the unobservable component, referred to as the *signal process* or *system process*, whereas Y_t is the observable part, called *observation process.* Thus information about X_t can only be obtained by extracting the information about X that is contained in the observation Y_t in a best possible way. In filter theory this is done by determining the conditional distribution of X_t given the information σ-field $\mathcal{F}_t^Y$ generated by $Y_s, 0 \le s \le t$. Or stated in an equivalent way, the objective is to compute the optimal filter as the conditional expectation

$$E_{\mathbb{P}}[f(X_t)|\mathcal{F}_t^Y]$$

for a rich enough class of functions f.

In the classical non-linear filter setting, the dynamics of the observation process Y_t is supposed to follow the following Itô process

$$dY_t = h(t, X_t)\, dt + dW_t\,,$$

where W_t is a Brownian motion independent of X. Under certain conditions on the drift $h(t, X_t)$ (see [KS], [K]), Kallianpur and Striebel derived a Bayes type formula for the conditional distribution expressed in terms of the so called unnormalized conditional distribution. In the special case when the dynamics of the signal follows an Itô diffusion

$$dY_t = b(t, X_t)\, dt + \sigma(t, X_t) dB_t\,,$$

for a second Brownian motion B_t, Zakai ([Z]) showed under certain conditions that the unormalized conditional density is the solution of an associated stochastic partial differential equation, the so called *Zakai equation.*

Date: August 30, 2007.
Key words and phrases. Non-linear Filtering, Gaussian noise, Cox noise, Bayes formula, Kallianpur-Striebel formula, Zakai euqation.

In this paper we extend the classical filter model to the following more general setting. For a general signal process X we suppose the observation model is given as

$$Y_t \;=\; \beta(t,X)+G_t+\int_0^t\int_{\mathbb{R}_0}\varsigma\,N_\lambda(dt,d\varsigma), \tag{1.1}$$

where

- G_t is a general Gaussian process with zero mean and continuous covariance function $R(s,t), 0\le s,t\le T$, that is independent of the signal process X.
- Let $\mathcal{F}_t^Y$ (respectively $\mathcal{F}_t^X$) denote the σ-algebra generated by $\{Y_s, 0\le s\le t\}$ (respectively $\{X_s, 0\le s\le t\}$) augmented by the null-sets. Define the filtration $(\mathcal{F}_t)_{0\le t\le T}$ through $\mathcal{F}_t := \mathcal{F}_T^X \vee \mathcal{F}_t^Y$. Then we assume that the process
$$L_t := \int_0^t\int_{\mathbb{R}_0}\varsigma\,N_\lambda(dt,d\varsigma)$$
is a pure jump $\mathcal{F}_t$-semimartingale determined through the integer valued random measure N_λ that has an $\mathcal{F}_t$-predictable compensator of the form
$$\mu(dt,d\varsigma,\omega)=\lambda(t,X,\varsigma)dt\nu(d\varsigma)$$
for a Lévy measure ν and a functional $\lambda(t,X(\omega),\varsigma)$. In particular, G_t and L_t are independent.
- The function $\beta:[0,T]\times\mathbb{R}^{[0,T]}\to\mathbb{R}$ is such that $\beta(t,\cdot)$ is $\mathcal{F}_t^X$-measurable and $\beta(\cdot,X(\omega))$ is in $H(R)$ for almost all ω, where $H(R)$ denotes the Hilbert space generated by $R(s,t)$ (see Section 2).

The observation dynamics consists thus of an information drift of the signal disturbed by some Gaussian noise plus a pure jump part whose jump intensity depends on the signal. Note that a jump process of the form given above is also referred to as *Cox process.*

The objective of the paper is in a first step to extend the Kallianpur-Striebel Bayes type formula to the generalized filter setting from above. When there are no jumps present in the observation dynamics (1.1) the corresponding formula has been developped in ([MM]). We will extend their way of reasoning to the situation including Cox noise.

In a second step we then derive a Zakai type measure valued stochastic differential equations for the unnormalized conditional distribution of the filter. For this purpose we assume the signal process X to be a Markov process with generator $\mathcal{O}_t := \mathcal{L}_t + \mathcal{B}_t$ given as

$$\begin{aligned}\mathcal{L}_t f(x) \;&:=\; b(t,x)\,\partial_x f(x)+\frac{1}{2}\sigma^2(t,x)\,\partial_{xx}f(x)\\ \mathcal{B}_t f(x) \;&:=\; \int_{\mathbb{R}_0}\{f(x+\gamma(t,x)\varsigma)-f(x)-\partial_x f(x)\gamma(t,x)\varsigma\}\,\upsilon(d\varsigma),\end{aligned}$$

with the coefficients $b(t,x),\sigma(t,x), and\gamma(t,x)$ and $f(x)$ being in $C_0^2(\mathbb{R})$ for every t. Here, $C_0^2(\mathbb{R})$ is the space of continuous functions with compact support and bounded derivatives up to order 2. Further, we develop a Zakai type stochastic parabolic integro partial differential equation for the unnormalized conditional density, given it exists. In the case the dynamics of X does not contain any jumps and the Gaussian noise G_t in the observation is Brownian motion, the corresponding Zakai equation was also studied in ([MP]). For further information on Zakai equations in a semimartingale setting we also refer to ([G1]) and ([G2]).

The remaining part of the paper is organized as follows. in Section 2 we briefly recall some theory of reproducing kernel Hilbert spaces. In Section 3 we obtain the Kallianpur-Striebel formula, before we discuss the Zakai type equations in Section 4.

2. Reproducing Kernel Hilbert Space and Stochastic Processes

A Hilbert space H consisting of real valued functions on some set $\mathbf{T}$ is said to be a *reproducing kernel Hilbert space* (RKHS), if there exists a function K on $\mathbf{T} \times \mathbf{T}$ with the following two properties: for every t in $\mathbf{T}$ and g in H,

(i) $K(\cdot, t) \in H$,
(ii) $(g(\cdot), K(\cdot, t)) = g(t)$. (The reproducing property)

K is called the *reproducing kernel* of H. The following basic properties can be found in [A].

(1) If a reproducing kernel exists, then it is unique.
(2) If K is the reproducing kernel of a Hilbert space H, then $\{K(\cdot, t), t \in \mathbf{T}\}$ spans H.
(3) If K is the reproducing kernel of a Hilbert space H, then it is nonnegative definite in the sense that for all $t_1, \ldots, t_n$ in $\mathbf{T}$ and $a_1, \ldots, a_n \in \mathbb{R}$

$$\sum_{i,j=1}^{n} K(t_i, t_j) a_i a_j \ \geq \ 0.$$

The converse of (3), stated in Theorem 2.1 below, is fundamental towards understanding the RKHS representation of Gaussian processes. A proof of the theorem can be found in [A].

Theorem 2.1 (E. H. Moore)**.** *A symmetric nonnegative definite function K on $\mathbf{T} \times \mathbf{T}$ generates a unique Hilbert space, which we denote by $H(K)$ or sometimes by $H(K, \mathbf{T})$, of which K is the reproducing kernel.*

Now suppose $K(s, t)$, $s, t \in \mathbf{T}$, is a nonnegative definite function. Then, by Theorem 2.1, there is a RKHS, $H(K, \mathbf{T})$, with K as its reproducing kernel. If we restrict K to $\mathbf{T}' \times \mathbf{T}'$ where $\mathbf{T}' \subset \mathbf{T}$, then K is still a nonnegative definite function. Hence K restricted to $\mathbf{T}' \times \mathbf{T}'$ will also correspond to a reproducing kernel Hilbert space $H(K, \mathbf{T}')$ of functions defined on $\mathbf{T}'$. The following result from ([A]; pp. 351) explains the relationship between these two.

Theorem 2.2. *Suppose $K_{\mathbf{T}}$, defined on $\mathbf{T} \times \mathbf{T}$, is the reproducing kernel of the Hilbert space $H(K_{\mathbf{T}})$ with the norm $\|\cdot\|$. Let $\mathbf{T}' \subset \mathbf{T}$, and $K_{\mathbf{T}'}$ be the restriction of $K_{\mathbf{T}}$ on $\mathbf{T}' \times \mathbf{T}'$. Then $H(K_{\mathbf{T}'})$ consists of all f in $H(K_{\mathbf{T}})$ restricted to $\mathbf{T}'$. Further, for such a restriction $f' \in H(K_{\mathbf{T}'})$ the norm $\|f'\|_{H(K_{\mathbf{T}'})}$ is the minimum of $\|f\|_{H(K_{\mathbf{T}})}$ for all $f \in H(K_{\mathbf{T}})$ whose restriction to $\mathbf{T}'$ is f'.*

If $K(s, t)$ is the covariance function for some zero mean process $Z_t, t \in \mathbf{T}$, then, by Theorem 2.1, there exists a unique RKHS, $H(K, \mathbf{T})$, for which K is the reproducing kernel. It is also easy to see (e.g., see Theorem 3D, [P1]) that there exists a congruence (linear, one-to-one, inner product preserving map) between $H(K)$ and $\overline{\mathrm{sp}}^{L^2}\{Z_t, t \in \mathbf{T}\}$ which takes

$K(\cdot,t)$ to Z_t. Let us denote by $\langle Z,h\rangle \in \overline{\text{sp}}^{L^2}\{Z_t, t\in \mathbf{T}\}$, the image of $h \in H(K,\mathbf{T})$ under the congruence.

We conclude the section with an important special case.

2.1. **A useful example.** Suppose the stochastic process Z_t is a Gaussian process given by

$$Z_t = \int_0^t F(t,u)dW_u,\ 0\le t\le T,$$

where $\int_0^t F^2(t,u)du < \infty$ for all $0 \le t \le T$ and W_u is Brownian motion. Then the covariance function

$$K(s,t) \equiv E(Z_sZ_t) = \int_0^{t\wedge s} F(t,u)F(s,u)du\,, \tag{2.1}$$

and the corresponding RKHS is given by

$$H(K) = \left\{ g : g(t) = \int_0^t F(t,u)g^*(u)du, 0 \le t \le T \right\} \tag{2.2}$$

for some (necessarily unique) $g^* \in \overline{\text{sp}}^{L^2}\{F(t,\cdot)1_{[0,t]}(\cdot), 0\le t\le T\}$, with the inner product

$$(g_1,g_2)_{H(K)} = \int_0^T g_1^*(u)g_2^*(u)du,$$

where

$$g_1(s) = \int_0^s F(s,u)g_1^*(u)du \quad \text{and} \quad g_2(s) = \int_0^s F(s,u)g_2^*(u)du.$$

For $0 \le t \le T$, by taking $K(\cdot,t)^*$ to be $F(t,\cdot)1_{[0,t]}(\cdot)$, we see, from (2.1) and (2.2), that $K(\cdot,t) \in H(K)$. To check the reproducing property suppose $h(t) = \int_0^t F(t,u)h^*(u)\,du \in H(K)$. Then

$$(h, K(\cdot,t))_{H(K)} = \int_0^T h^*(u)K(\cdot,t)^*\,du = \int_0^t h^*(u)F(t,u)\,du = h(t).$$

Also, in this case, it is very easy to check (cf. [P2], Theorem 4D) that the congruence between $H(K)$ and $\overline{\text{sp}}^{L^2}\{Z_t, t\in\mathbf{T}\}$ is given by

$$\langle Z,g\rangle \ = \int_0^T g^*(u)dW_u. \tag{2.3}$$

3. The Filter Setting and a Bayes Formula

Assume a partially observable process $(X,Y) = (X_t,Y_t)_{0\le t\le T} \in \mathbb{R}^2$ defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. The real valued process X_t stands for the unobservable component, referred to as the *signal process*, whereas Y_t is the observable part, called *observation process*. In particular, we assume that the dynamics of the observation process is given as follows:

$$Y_t \ = \ \beta(t,X) + G_t + \int_0^t \int_{\mathbb{R}_0} \varsigma\, N_\lambda(dt,d\varsigma), \tag{3.1}$$

where

- G_t is a Gaussian process with zero mean and continuous covariance function $R(s,t), 0 \leq s,t \leq T$, that is independent of the signal process X.
- The function $\beta : [0,T] \times \mathbb{R}^{[0,T]} \to \mathbb{R}$ is such that $\beta(t,\cdot)$ is $\mathcal{F}_t^X$-measurable and $\beta(\cdot, X(\omega))$ is in $H(R)$ for almost all ω, where $H(R)$ denotes the Hilbert space generated by $R(s,t)$ (see Section 2).
- Let $\mathcal{F}_t^Y$ (respectively $\mathcal{F}_t^X$) denote the σ-algebra generated by $\{Y_s, 0 \leq s \leq t\}$ (respectively $\{X_s, 0 \leq s \leq t\}$) augmented by the null-sets. Define the filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ through $\mathcal{F}_t := \mathcal{F}_T^X \vee \mathcal{F}_t^Y$. Then we assume that the process
$$L_t := \int_0^t \int_{\mathbb{R}_0} \varsigma\, N_\lambda(dt, d\varsigma)$$
is a pure jump $\mathcal{F}_t$-semimartingale determined through the integer valued random measure N_λ that has an $\mathcal{F}_t$-predictable compensator of the form
$$\mu(dt, d\varsigma, \omega) = \lambda(t, X, \varsigma) dt \nu(d\varsigma)$$
for a Lévy measure ν and a functional $\lambda(t, X(\omega), \varsigma)$.
- The functional $\lambda(t, X, \varsigma)$ is assumed to be strictly positive and such that
$$\begin{aligned} &\int_0^T \int_{\mathbb{R}_0} \log^2\left(\lambda(s,X,\varsigma)\right)\, \mu(ds, d\varsigma) < \infty \ \text{ a.s.} \\ &\int_0^T \int_{\mathbb{R}_0} \log^2\left(\lambda(s,X,\varsigma)\right)\, ds\, \nu(d\varsigma) < \infty \ \text{ a.s.} \end{aligned} \tag{3.2}$$
and
$$\begin{aligned} \Lambda_t \ &:= \ \exp\left\{ \int_0^t \int_{\mathbb{R}_0} \log\left(\frac{1}{\lambda(s,X,\varsigma)}\right) \widetilde{N}_\lambda(ds, d\varsigma) \right. \\ &\left. + \int_0^t \int_{\mathbb{R}_0} \left(\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right) - \frac{1}{\lambda(s,X,\varsigma)} + 1 \right) \mu(ds, d\varsigma) \right\} \end{aligned} \tag{3.3}$$
is a well defined $\mathcal{F}_t$-martingale. Here $\widetilde{N}_\lambda(ds, d\varsigma)$ stands for the compensated jump measure
$$\widetilde{N}_\lambda(ds, d\varsigma) := N_\lambda(ds, d\varsigma) - \mu(dt, d\varsigma).$$

Remark 3.1. Note that the specific from of the predictable compensator $\mu(dt, d\varsigma, \omega)$ implies that L_t is a process with conditionally independent increments with respect to the σ-algebra $\mathcal{F}_T^X$, i.e.
$$E_{\mathbb{P}}[f(L_t - L_s) 1_A | \mathcal{F}_T^X] = E_{\mathbb{P}}[f(L_t - L_s) | \mathcal{F}_T^X]\, E_{\mathbb{P}}[1_A | \mathcal{F}_T^X],$$
for all bounded measurable functions f, $A \in \mathcal{F}_s$, and $0 \leq s < t \leq T$ (see for example Th. 6.6 in [JS]). Also, it follows that the processes G is independent from the random measure $N_\lambda(ds, d\varsigma)$.

Given a Borel measurable function f, our non-linear filtering problem then comes down to determine the least square estimate of $f(X_t)$, given the observations up to time t. In other words, the problem consists in evaluating the *optimal filter*
$$E_{\mathbb{P}}[f(X_t) \,|\mathcal{F}_t^Y] \,. \tag{3.4}$$

In this section we want to derive a Bayes formula for the optimal filter (3.4) by an extension of the reference measure method presented in [MM] for the purely Gaussian case. For this purpose, define for each $0 \le t \le T$ with $\beta(\cdot) = \beta(\cdot, X)$

$$\Lambda'_t := \exp\left\{ -\langle G, \beta\rangle_t - \frac{1}{2}\|\beta\|_t^2 \right\}$$

Then the main tool is the following extension of Theorem 3.1 in [MM]

Lemma 3.2. *Define*

$$d\mathbb{Q} := \Lambda_t \Lambda'_t \, d\mathbb{P}.$$

Then $\mathbb{Q}_t$ is a probability measure, and under $\mathbb{Q}_t$ we have that

$$Y_t = \widetilde{G}_t + L_t,$$

where $\widetilde{G}_s = \beta(s, X) + G_s$, $0 \le s \le t$, is a Gaussian process with zero mean and covariance function R, L_s, $0 \le s \le t$, is a pure jump Lévy process with Lévy measure ν, and the process X_s, $0 \le s \le T$ has the same distribution as under $\mathbb{P}$. Further, the processes $\widetilde{G}$, L and X are independent under $\mathbb{Q}_t$.

Proof. Fix $0 \le t \le T$. First note that since $\beta(\cdot) \in H(R)$ almost surely, we have by Theorem 2.2 that $\beta|_{[0,t]} \in H(R; t)$ almost surely. Further, by the independence of the Gaussian process G from X and from the random measure $N_\lambda(ds, d\varsigma)$ it follows that

$$E_{\mathbb{P}}[\Lambda_t \Lambda'_t] = E_{\mathbb{P}}[E_{\mathbb{P}}[\Lambda_t | \mathcal{F}_T^X] \, E_{\mathbb{P}}[\Lambda'_t | \mathcal{F}_T^X]].$$

Since for $f \in H(R; t)$ the random variable $\langle G, f\rangle_t$ is Gaussian with zero mean and variance $\|f\|_t^2$, it follows again by the independence of G from X and the martingale property of Λ_t that $E_{\mathbb{P}}[\Lambda_t \Lambda'_t] = 1$, and $\mathbb{Q}_t$ is a probability measure.

Now take $0 \le s_1, ..., s_m \le t$, $0 \le r_1, ..., r_p \le t$, $0 \le t_1, ..., t_n \le T$ and real numbers $\lambda_1, ..., \lambda_m$, $\gamma_1, ..., \gamma_p$, $\alpha_1, ..., \alpha_n$ and consider the joint characteristic function

$$\begin{aligned}
&E_{\mathbb{Q}_t}\left[e^{i\sum_{j=1}^n \alpha_j X_{t_j} + i\sum_{i=1}^m \lambda_i \widetilde{G}_{s_i} + i\sum_{k=1}^p \gamma_k (L_{r_k} - L_{r_{k-1}})} \right] \\
&= E_{\mathbb{P}}\left[e^{i\sum_{j=1}^n \alpha_j X_{t_j} + i\sum_{i=1}^m \lambda_i \widetilde{G}_{s_i} + i\sum_{k=1}^p \gamma_k (L_{r_k} - L_{r_{k-1}})} \Lambda_t \Lambda'_t \right] \\
&= E_{\mathbb{P}}\left[e^{i\sum_{j=1}^n \alpha_j X_{t_j}} E_{\mathbb{P}}[e^{i\sum_{i=1}^m \lambda_i \widetilde{G}_{s_i}} \Lambda'_t | \mathcal{F}_T^X] \, E_{\mathbb{P}}[e^{i\sum_{k=1}^p \gamma_k (L_{r_k} - L_{r_{k-1}})} \Lambda_t \, | \mathcal{F}_T^X] \right].
\end{aligned}$$

Here, for computational convenience, the part of the characteristic function that concerns L is formulated in terms of increments of L (where we set $r_0 = 0$). Now, as in Theorem 3.1 in [MM], we get by the independence of G from X that

$$E_{\mathbb{P}}[e^{i\sum_{i=1}^m \lambda_i \widetilde{G}_{s_i}} \Lambda'_t | \mathcal{F}_T^X] = e^{-\sum_{i,l=1}^m \lambda_i \lambda_l R(s_i, s_l)},$$

which is the characteristic function of a Gaussian process with mean zero and covariance function R.

Further, by the conditional independent increments of L we get like in the proof of Th. 6.6 in [JS] that

$$E_{\mathbb{P}}\left[e^{\int_r^u \int_{\mathbb{R}_0} \delta(s,X,\varsigma) \, \widetilde{N}_\lambda(ds, d\varsigma)} | \mathcal{F}_T^X \right] = e^{\int_r^u \int_{\mathbb{R}_0} \left(e^{\delta(s,X,\varsigma)} - 1 - \delta(s,X,\varsigma) \right) \mu(dt, d\varsigma)}$$

for $0 \leq r \leq u \leq T$. So that for one increment one obtains

$$
\begin{aligned}
&E_{\mathbb{P}}\left[e^{i\gamma(L_u-L_r)}\Lambda_t\,|\mathcal{F}_T^X\right] \\
&\quad = E_{\mathbb{P}}\left[\exp\left\{\int_r^u\int_{\mathbb{R}_0}\left(i\gamma\varsigma+\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right)\right)\widetilde{N}_\lambda(ds,d\varsigma)\right.\right. \\
&\qquad\quad \left.\left.+\int_r^u\int_{\mathbb{R}_0}\left(i\gamma\varsigma+\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right)-\frac{1}{\lambda(s,X,\varsigma)}+1\right)\mu(dt,d\varsigma)\right\}|\mathcal{F}_T^X\right] \\
&\quad = E_{\mathbb{P}}\left[\exp\left\{\int_r^u\int_{\mathbb{R}_0}\left(e^{i\gamma\varsigma+\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right)}-1-i\gamma\varsigma-\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right)\right)\mu(dt,d\varsigma)\right.\right. \\
&\qquad\quad \left.\left.+\int_r^u\int_{\mathbb{R}_0}\left(i\gamma\varsigma+\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right)-\frac{1}{\lambda(s,X,\varsigma)}+1\right)\mu(dt,d\varsigma)\right\}|\mathcal{F}_T^X\right] \\
&\quad = E_{\mathbb{P}}\left[\exp\left\{\int_r^u\int_{\mathbb{R}_0}\left(e^{i\gamma\varsigma+\log\left(\frac{1}{\lambda(s,X,\varsigma)}\right)}-\frac{1}{\lambda(s,X,\varsigma)}\right)\lambda(t,X,\varsigma)dt\nu(d\varsigma)\right\}|\mathcal{F}_T^X\right] \\
&\quad = \exp\left\{(u-r)\int_{\mathbb{R}_0}\left(e^{i\gamma\varsigma}-1\right)\nu(d\varsigma)\right\}.
\end{aligned}
$$

The generalization to the sum of increments is straightforward and one obtains the characteristic function of the finite dimensional distribution of a Lévy process (of finite variation):

$$
E_{\mathbb{P}}[e^{i\sum_{k=1}^p\gamma_k(L_{r_k}-L_{r_{k-1}})}\Lambda_t|\mathcal{F}_T^X] = \exp\left\{\sum_{k=1}^p(r_k-r_{k-1})\int_{\mathbb{R}_0}\left(e^{i\gamma_k\varsigma}-1\right)\nu(d\varsigma)\right\}.
$$

All together we end up with

$$
\begin{aligned}
&E_{\mathbb{Q}_t}\left[e^{i\sum_{j=1}^n\alpha_jX_{t_j}+i\sum_{i=1}^m\lambda_i\widetilde{G}_{s_i}+i\sum_{k=1}^p\gamma_k(L_{r_k}-L_{r_{k-1}})}\right] \\
&\quad = E_{\mathbb{P}}\left[e^{i\sum_{j=1}^n\alpha_jX_{t_j}}\right]\cdot e^{-\sum_{i,l=1}^m\lambda_i\lambda_lR(s_i,s_l)}\cdot e^{\sum_{k=1}^p(r_k-r_{k-1})\int_{\mathbb{R}_0}\left(e^{i\gamma_k\varsigma}-1\right)\nu(d\varsigma)},
\end{aligned}
$$

which completes the proof. □

Remark 3.3. Note that in case G is Brownian motion Lemma 3.2 is just the usual Girsanov theorem for Brownian motion and random measures. In this case, it follows from Cameron-Martin's result and the fact that X is independent of G that $\Lambda_t\Lambda_t'$ is a martingale and $d\mathbb{Q}$ is a probability measure.

Now, the inverse Radon-Nikodym derivative

$$
\frac{d\mathbb{P}}{d\mathbb{Q}_t} = (\Lambda_t)^{-1}(\Lambda_t')^{-1}
$$

is $\mathbb{Q}_t$-a.s. by condition (3.2) and an argument like in ([MM], p. 857) given through

$$
\begin{aligned}
(\Lambda_t)^{-1} &= \exp\left\{\int_0^t\int_{\mathbb{R}_0}\log\left(\lambda(s,X,\varsigma)\right)\widetilde{N}(ds,d\varsigma)\right. \\
&\qquad \left.+\int_0^t\int_{\mathbb{R}_0}\left(\log\left(\lambda(s,X,\varsigma)\right)-\lambda(s,X,\varsigma)+1\right)ds\,\nu(d\varsigma)\right\} \\
(\Lambda_t')^{-1} &= \exp\left\{\langle\widetilde{G},\beta\rangle_t-\frac{1}{2}\|\beta\|_t^2\right\}.
\end{aligned}
$$

Here

$$\widetilde{N}(ds,d\varsigma) := N_\lambda(ds,d\varsigma) - dt\nu(d\varsigma)$$

is now a compesnsated Poisson random measure under $\mathbb{Q}_t$. Then we have by the Bayes formula for conditional expectation for any $\mathcal{F}_T^X$-measurable integrable function $g(T,X)$

$$E_{\mathbb{P}}\left[g(T,X)\,|\mathcal{F}_t^Y\right] = \frac{E_{\mathbb{Q}_t}\left[g(T,X)(\Lambda_t)^{-1}(\Lambda_t')^{-1}\,|\mathcal{F}_t^Y\right]}{E_{\mathbb{Q}_t}\left[(\Lambda_t)^{-1}(\Lambda_t')^{-1}\,|\mathcal{F}_t^Y\right]}.$$

From Lemma 3.2 we know that the processes $\left(\widetilde{G}_s\right)_{0\le s\le t}$, $(L_s)_{0\le s\le t}$, and $(X_s)_{0\le s\le T}$ are independent under $\mathbb{Q}_t$ and that the distribution of X is the same under $\mathbb{Q}_t$ as under $\mathbb{P}$. Hence conditional expectations of the form $E_{\mathbb{Q}_t}\left[\phi(X,\widetilde{G},L)\,|\mathcal{F}_t^Y\right]$ can be computed as

$$\begin{aligned} E_{\mathbb{Q}_t}\left[\phi(X,\widetilde{G},L)\,|\mathcal{F}_t^Y\right](\omega) &= \int_\Omega \phi(X(\hat\omega),\widetilde{G}(\omega),L(\omega))\,\mathbb{Q}_t(d\hat\omega) \\ = \int_\Omega \phi(X(\hat\omega),\widetilde{G}(\omega),L(\omega))\,\mathbb{P}(d\hat\omega) &= E_{\hat{\mathbb{P}}}\left[\phi(X(\hat\omega),\widetilde{G}(\omega),L(\omega))\right]. \end{aligned}$$

where $(\omega,\hat\omega)\in\Omega\times\Omega$ and the index $\hat{\mathbb{P}}$ denotes integration with respect to $\hat\omega$. Consequently, we get the following Bayes formula for the optimal filter

Theorem 3.4. *Under the above specified conditions, for any $\mathcal{F}_T^X$-measurable integrable function $g(T,X)$*

$$\begin{aligned} E_{\mathbb{P}}\left[g(T,X)\,|\mathcal{F}_t^Y\right] &= \frac{\int_\Omega g(T,X(\hat\omega))\alpha_t(\omega,\hat\omega)\alpha_t'(\omega,\hat\omega)\,\mathbb{P}(d\hat\omega)}{\int_\Omega \alpha_t(\omega,\hat\omega)\alpha_t'(\omega,\hat\omega)\,\mathbb{P}(d\hat\omega)} \\ &= \frac{E_{\hat{\mathbb{P}}}\left[g(T,X(\hat\omega))\alpha_t(\omega,\hat\omega)\alpha_t'(\omega,\hat\omega)\right]}{E_{\hat{\mathbb{P}}}\left[\alpha_t(\omega,\hat\omega)\alpha_t'(\omega,\hat\omega)\right]}, \end{aligned}$$

where

$$\begin{aligned} \alpha_t(\omega,\hat\omega) &= \exp\left\{\int_0^t\int_{\mathbb{R}_0}\log\left(\lambda(s,X(\hat\omega),\varsigma)\right)\widetilde{N}(\omega,ds,d\varsigma)\right. \\ &\qquad \left. + \int_0^t\int_{\mathbb{R}_0}\left(\log\left(\lambda(s,X(\hat\omega),\varsigma)\right) - \lambda(s,X(\hat\omega),\varsigma)+1\right)ds\,\nu(d\varsigma)\right\} \\ \alpha_t'(\omega,\hat\omega) &= \exp\left\{\langle\widetilde{G}(\omega),\beta(\cdot,\hat\omega)\rangle_t - \frac{1}{2}\|\beta(\cdot,\hat\omega)\|_t^2\right\}. \end{aligned}$$

4. Zakai type equations

Using the Bayes formula from above we now want to proceed further in deriving a Zakai type equations for the unnormalized filter. This equation is basic in order to obtain the filter recursively. To this end we have to impose certain restrictions on both the signal process and the Gaussian part of the observation process.

Regarding the signal process X, we assume its dynamics to be Markov. More precisely, we consider the parabolic integro-differential operator $\mathcal{O}_t := \mathcal{L}_t + \mathcal{B}_t$, where

$$\mathcal{L}_t f(x) \;:=\; b(t,x)\,\partial_x f(x) + \frac{1}{2}\sigma^2(t,x)\,\partial_{xx} f(x)$$

$$\mathcal{B}_t f(x) \;:=\; \int_{\mathbb{R}_0} \{f(x+\gamma(t,x)\varsigma) - f(x) - \partial_x f(x)\gamma(t,x)\varsigma\}\, \upsilon(d\varsigma),$$

for $f \in C_0^2(\mathbb{R})$. Here, $C_0^2(\mathbb{R})$ is the space of continuous functions with compact support and bounded derivatives up to order 2. Further, we suppose that $b(t,\cdot)$, $\sigma(t,\cdot)$, and $\gamma(t,\cdot)$ are in $C_0^2(\mathbb{R})$ for every t and that $\upsilon(d\varsigma)$ is a Lévy measure with second moment. The signal process X_t, $0 \le t \le T$, is then assumed to be a solution of the martingale problem corresponding to $\mathcal{O}_t$, i.e.

$$f(X_t) - \int_0^t (\mathcal{O}_u f)(X_u)\, du$$

is a $\mathcal{F}_t^X$-martingale with respect to $\mathbb{P}$ for every $f \in C_0^2(\mathbb{R})$.

Further, we restrict the Gaussian process G of the observation process in (3.1) to belong to the special case presented in Section 2.1, i.e.

$$G_t = \int_0^t F(t,s)\, dW_s,$$

where W_t is Brownian motion and $F(t,s)$ is a deterministic function such that $\int_0^t F^2(t,s)\, ds$, $0 \le t \le T$. Note that this type of processes both includes Ornstein-Uhlenbeck processes as well as fractional Brownian motion. Then $\beta(t,X)$ will be of the form

$$\beta(t,X) = \int_0^t F(t,s)h(s,X_s)\, ds.$$

Further, with

$$\widetilde{W}_t := \int_0^t h(s,X_s)\, ds + W_t$$

we get $\langle \widetilde{G}, \beta\rangle_t = \int_0^t h(s,X_s)\, d\widetilde{W}_s$ and $\|\beta\|_t^2 = \int_0^t h^2(s,X_s)\, ds$, and $\alpha_t'(\omega,\hat\omega)$ in Theorem 3.4 becomes

$$\alpha_t'(\omega,\hat\omega) = \exp\left\{ \int_0^t h(s,X_s(\hat\omega))\, d\widetilde{W}_s(\omega) - \frac{1}{2}\int_0^t h^2(s,X_s(\hat\omega))\, ds \right\}.$$

Note that in this case $\widetilde{W}_s$, $0 \le s \le t$, is a Brownian motion under $\mathbb{Q}_t$.

For $f \in C_0^2(\mathbb{R})$ we now define the unnormalized filter $V_t(f) = V_t(f)(\omega)$ by

$$V_t(f)(\omega) := \int_\Omega f(X_t(\hat\omega))\alpha_t(\omega,\hat\omega)\alpha_t'(\omega,\hat\omega)\, \mathbb{P}(d\hat\omega) = E_{\hat{\mathbb{P}}}\left[f(X_t(\hat\omega))\alpha_t(\omega,\hat\omega)\alpha_t'(\omega,\hat\omega)\right].$$

Then this unnormalized filter obeys the following dynamics

Theorem 4.1. *(Zakai equation I) Under the above specified assumptions, the unnormalized filter $V_t(f)$ satisfies the equation*

$$dV_t\Big(f(\cdot)\Big)(\omega) \;=\; V_t\Big(\mathcal{O}_t f(\cdot)\Big)(\omega)\, dt + V_t\Big(h(t,\cdot)f(\cdot)\Big)(\omega)\, d\widetilde{W}_t(\omega) \tag{4.1}$$

$$+ \int_{\mathbb{R}_0} V_t\Big(\,(\lambda(t,\cdot,\varsigma)-1)\, f(\cdot)\Big)(\omega)\, \widetilde{N}(\omega,dt,d\varsigma). \tag{4.2}$$

Proof. Set

$$g_t(\hat{\omega}) := f(X_T(\hat{\omega})) - \int_t^T (\mathcal{O}_s f)(X_s(\hat{\omega}))\, ds.$$

Then, by our assumptions on the coefficients b, σ, γ and on the Lévy measure $\upsilon(d\varsigma)$, we have $|g_t| < C$ for some constant C. Since $f(X_t) - \int_0^t \mathcal{O}_t f(X_s)\, ds$ is a martingale we obtain

$$E_{\hat{\mathbb{P}}}\left[g_t \,|\mathcal{F}_t^{X(\hat{\omega})}\right] = f(X_t), \qquad 0 \le t \le T. \tag{4.3}$$

If we denote

$$\Gamma_t(\omega,\hat{\omega}) := \alpha_t(\omega,\hat{\omega})\alpha_t'(\omega,\hat{\omega}),$$

then, because $\Gamma_t(\omega,\hat{\omega})$ is $\mathcal{F}_t^{X(\hat{\omega})}$-measurable for each ω, equation (4.3) implies that

$$\begin{aligned} V_t(f) &= E_{\hat{\mathbb{P}}}\left[f(X_t(\hat{\omega}))\Gamma_t(\omega,\hat{\omega})\right] \\ &= E_{\hat{\mathbb{P}}}\left[E_{\hat{\mathbb{P}}}\left[g_t(\hat{\omega})\Gamma_t(\omega,\hat{\omega}) \,|\mathcal{F}_t^{X(\hat{\omega})}\right]\right] \\ &= E_{\hat{\mathbb{P}}}\left[g_t(\hat{\omega})\Gamma_t(\omega,\hat{\omega})\right] \end{aligned}$$

By definition of g_t,

$$dg_t(\hat{\omega}) = (\mathcal{O}_t f)(X_t(\hat{\omega}))\, dt.$$

Also, $\Gamma_t = \Gamma_t(\omega,\hat{\omega})$ is the Doléans-Dade solution of the following linear SDE

$$d\Gamma_t = h(t, X_t(\hat{\omega})\Gamma_t\, d\widetilde{W}_t(\omega) + \int_{\mathbb{R}_0} \left(\lambda(t, X_t(\hat{\omega}),\varsigma) - 1\right) \Gamma_t\, \widetilde{N}(\omega, dt, d\varsigma).$$

So we get

$$\begin{aligned} E_{\hat{\mathbb{P}}}\left[g_t(\hat{\omega})\Gamma_t\right] &= E_{\hat{\mathbb{P}}}\left[g_0(\hat{\omega})\Gamma_0\right] + E_{\hat{\mathbb{P}}}\left[\int_0^t (\mathcal{O}_s f)(X_s(\hat{\omega}))\Gamma_s\, ds\right] \\ &\quad + E_{\hat{\mathbb{P}}}\left[\int_0^t h(s, X_s(\hat{\omega}) g_s(\hat{\omega})\Gamma_s\, d\widetilde{W}_s(\omega)\right] \\ &\quad + E_{\hat{\mathbb{P}}}\left[\int_0^t \int_{\mathbb{R}_0} \left(\lambda(s, X_s(\hat{\omega}),\varsigma) - 1\right) g_s(\hat{\omega})\Gamma_s\, \widetilde{N}(\omega, ds, d\varsigma)\right]. \end{aligned}$$

The first term on the right hand side equals $f(X_0)$, and for the second one we can invoke Fubini's theorem to get

$$E_{\hat{\mathbb{P}}}\left[\int_0^t (\mathcal{O}_s f)(X_s(\hat{\omega}))\Gamma_s\, ds\right] = \int_0^t E_{\hat{\mathbb{P}}}\left[(\mathcal{O}_s f)(X_s(\hat{\omega}))\Gamma_s\right]\, ds = \int_0^t V_s\Big(\mathcal{O}_s f(\cdot)\Big)(\omega)\, ds.$$

For the third term we employ the stochastic Fubini theorem for Brownian motion (see for example 5.14 in [LS]) in order to get

$$\begin{aligned} &E_{\hat{\mathbb{P}}}\left[\int_0^t h(s, X_s(\hat{\omega}) g_s(\hat{\omega})\Gamma_s\, d\widetilde{W}_s(\omega)\right] \\ &= \int_0^t E_{\hat{\mathbb{P}}}\left[h(s, X_s(\hat{\omega}) g_s(\hat{\omega})\Gamma_s\right]\, d\widetilde{W}_s(\omega) \\ &= \int_0^t E_{\hat{\mathbb{P}}}\left[h(s, X_s(\hat{\omega})\Gamma_s E_{\hat{\mathbb{P}}}\left[g_s(\hat{\omega}) \,|\mathcal{F}_s^{X(\hat{\omega})}\right]\right]\, d\widetilde{W}_s(\omega) \end{aligned}$$

$$= \int_0^t E_{\hat{\mathbb{P}}} \left[h(s, X_s(\hat{\omega}) \Gamma_s f(X_s(\hat{\omega})) \right] d\widetilde{W}_s(\omega)$$
$$= \int_0^t V_s \Big(h(s,\cdot) f(\cdot) \Big)(\omega)\, d\widetilde{W}_s(\omega).$$

Further, one easily sees that the analogue stochastic Fubini theorem for compensated Poisson random measures holds, and we get analogously for the last term

$$E_{\hat{\mathbb{P}}} \left[\int_0^t \int_{\mathbb{R}_0} (\lambda(s, X_s(\hat{\omega}), \varsigma) - 1)\, g_s(\hat{\omega}) \Gamma_s\, \widetilde{N}(\omega, ds, d\varsigma) \right]$$
$$= \int_0^t \int_{\mathbb{R}_0} V_s \Big(\, (\lambda(s, \cdot, \varsigma) - 1)\, f(\cdot) \Big)(\omega)\, \widetilde{N}(\omega, ds, d\varsigma),$$

which completes the proof. □

If one further assume that the filter has a so called unormalized conditional density $u(t,x)$ then we can derive a stochastic integro-PDE determining $u(t,x)$ which for the Brownian motion case was first established in [Z] and usually is referred to as Zakai equation.

Definition 4.2. *We say that a process $u(t,x) = u(\omega, t, x)$ is the unnormalized conditional density of the filter if*

$$V_t(f)(\omega) = \int_{\mathbb{R}} f(x) u(\omega, t, x)\, dx \tag{4.4}$$

for all bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$.

From now on we restrict the intergo part $\mathcal{B}_t$ of the operator $\mathcal{O}_t$ to be the one of a pure jump Lévy process, i.e. $\gamma = 1$, and we assume the initial value $X_0(\omega)$ of the signal process to posses a density denoted by $\xi(x)$. Then the following holds

Theorem 4.3. *(Zakai equation II) Suppose the unnormalized conditional density $u(t,x)$ of our filter exists. Then, provided a solution exists, $u(t,x)$ solves the following stochastic integro-PDE*

$$\begin{aligned} du(t,x) &= \mathcal{O}_t^* u(t,x) dt + h(t,x) u(t,x) d\widetilde{W}_t(\omega) \\ &\qquad + \int_{\mathbb{R}_0} (\lambda(t,x,\varsigma) - 1) u(t,x) \widetilde{N}(\omega, dt, d\varsigma) \\ u(0,x) &= \xi(x). \end{aligned} \tag{4.5}$$

Here $\mathcal{O}_t^* := \mathcal{L}_t^* + \mathcal{B}_t^*$ is the adjoint operator of $\mathcal{O}_t$ given through

$$\begin{aligned} \mathcal{L}_t^* f(x) &:= -\partial_x \left(b(t,x) f(x) \right) + \frac{1}{2} \partial_{xx} \left(\sigma^2(t,x) f(x) \right) \\ \mathcal{B}_t^* f(x) &:= \int_{\mathbb{R}_0} \left\{ f(x - \varsigma) - f(x) + \partial_x f(x) \varsigma \right\} \upsilon(d\varsigma), \end{aligned}$$

for $f \in C_0^2(\mathbb{R})$.

For sufficient conditions on the coefficients under which there exists a classical solution of (4.5) see for example [MP]; in [M] the existence of solutions in a generalized sense of stochastic distributions is treated.

Proof. By (4.1) and (4.4) we have for all $f \in C_0^\infty(\mathbb{R})$

$$\begin{aligned}\int_{\mathbb{R}} f(x)u(t,x)\,dx &= \int_{\mathbb{R}} f(x)\xi(x)\,dx + \int_0^t \int_{\mathbb{R}} u(s,x)\mathcal{O}_s^* f(x)\,dxds \\ &\quad + \int_0^t \int_{\mathbb{R}} u(s,x)h(s,x)f(x)\,dxd\widetilde{W}_s(\omega) \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \int_{\mathbb{R}} u(s,x)\left(\lambda(s,x,\varsigma)-1\right)f(x)\,dx\widetilde{N}(\omega,ds,d\varsigma)\end{aligned}$$

Now, using integration by parts, we get

$$\int_{\mathbb{R}} u(s,x)\mathcal{L}_s^* f(x)\,dx = \int_{\mathbb{R}} f(x)\mathcal{O}_s^* u(s,x)\,dx. \tag{4.6}$$

Further it holds again integration by parts and by substitution that

$$\int_{\mathbb{R}} u(s,x)\mathcal{B}_s^* f(x)\,dx = \int_{\mathbb{R}} f(x)\mathcal{B}_s^* u(s,x)\,dx. \tag{4.7}$$

Fubini together with (4.6) and (4.7) then yields

$$\begin{aligned}\int_{\mathbb{R}} f(x)u(t,x)\,dx &= \int_{\mathbb{R}} f(x)\xi(x)\,dx + \int_{\mathbb{R}} f(x)\left(\int_0^t \mathcal{O}_s^* u(s,x)\,ds\right)dx \\ &\quad + \int_{\mathbb{R}} f(x)\left(\int_0^t u(s,x)h(s,x)\,d\widetilde{W}_s(\omega)\right)dx \\ &\quad + \int_{\mathbb{R}} f(x)\left(\int_0^t \int_{\mathbb{R}_0} u(s,x)\left(\lambda(s,x,\varsigma)-1\right)\widetilde{N}(\omega,ds,d\varsigma)\right)dx\end{aligned}$$

Since this holds for all $f \in C_0^\infty(\mathbb{R})$ we get (4.5). □

Acknowledgements: This work was started when V. Mandrekar was at the CMA,University of Oslo. He thanks the CMA for support and Prof. Øksendal for interesting discussions and hospitality.

References

[A] N. Aronszajn (1950): *Theory of reproducing kernels,* Trans. Amer. Math. Soc., 68, pp. 337-404.

[G1] Grigelionis B.: Stochastic non-linear filtering equations and semimartingales. In Lecture Notes Math. 972, eds Mitter, S.K., Moro, A., Springer-Verlag, Berlin (1982).

[G2] Grigelionis B.: Stochastic evolution equations and densities of the conditional distributions,Lecture notes in Control and Info. Sci., 1983, p. 49.

[JS] Jacod, J., Shiryaev, A.N. (2002): *Limit Theorems for Stochastic Processes*. Springer, 2. ed., Berlin Heidelberg New York.

[K] G. Kallianpur (1980): *Stochastic Filtering Theory*. Springer Verlag, New York, Heidelberg, Berlin, 1980.

[KS] G. Kallianpur, C. Striebel (1968): *Estimations of stochastic processes : Arbitrary system process with additive white noise observation errors*. Ann. Math. Statis., 39 (1968), pp. 785-801.

[LS] Liptser R.S., Shiryaev A.N. (1977): *Statistics of Random Processes*. Vol. 1, Springer-Verlag, NY, 1977.

[MM] P. Mandal and V. Mandrekar (2000): *A Bayes formula for Gaussian noise processes and its applications*, SIAM J. of Control and Optimization 39 852871.

[M] Meyer-Brandis T. (2005): *Stochastic Feynman-Kac Equations associated to Lévy-Itô Diffusions*, Stochastic Analysis and Applications, to appear.

[MP] Meyer-Brandis, T., Proske, F.: Explicit Solution of a Non-Linear Filtering Problem for Lévy Processes with Application to Finance, Appl. Math. Optim. 50 (2004), pp.119-134.

[P1] E. Parzen(1961): *Regression Analysis of Continuous Parameter Time Series,* in Proc. 4th Berkeley Sympos. Math. Statis. and Prob., Vol. I, Univ. California Press, Berkeley, Calif., 1961, pp. 469-489.

[P2] E. Parzen(1967): *Statistical inference on time series by Hilbert space methods, I,* in Time Series Analysis Papers, Holden-Day, San Francisco, Cambridge, London, Amsterdam, 1967, pp. 251-382.

[Z] Zakai, M. (1969): *On the optimal filtering of diffusion processes.* Z. Wahrsch. verw. Geb., No. 11, 230-243, (1969).

(Vidyadhar Mandrekar), Department of Statistics and Probability, Michigan State University, East Lansing, MI48824, USA

(Thilo Meyer-Brandis), Center of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway,

(Frank Proske), Center of Mathematics for Applications, University of Oslo, P.O. Box 1053, Blindern, N–0316 Oslo, Norway,